Order Code RL32536

The Multi-State Anti-Terrorism Information Exchange (MATRIX) Pilot Project

August 18, 2004

William J. Krouse
Analyst in Social Legislation
Domestic Social Policy Division

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
18 AUG 2004	**N/A**	**-**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
The Multi-State Anti-Terrorism Information Exchange (MATRIX) Pilot Project	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Congressional Research Service Library of Congress 101 Independence Avenue, SE Washington, D.C. 20540-7500**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release, distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			
unclassified	**unclassified**	**unclassified**	**SAR**	**13**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

The Multi-State Anti-Terrorism Information Exchange (MATRIX) Pilot Project

Summary

This report provides an overview of the Multi-State Anti-Terrorism Information Exchange (MATRIX) pilot project, which leverages advanced computer/information management capabilities to more quickly access, share, and analyze *public* records to help law enforcement generate leads, expedite investigations, and possibly prevent terrorist attacks. The pilot project is intended to demonstrate the effective use of such capabilities, but it is less clear whether the project has been designed to prevent unnecessary intrusions on privacy.

The MATRIX pilot project is being administered by the Institute for Intergovernmental Research for the Department of Homeland Security (DHS). Project security and access to the MATRIX system is managed by the Florida Department of Law Enforcement. The project is being funded by the DHS Office of Domestic Preparedness ($8 million) and the Department of Justice's Bureau of Justice Assistance ($4 million).

Privacy advocates, civil libertarians, and others oppose MATRIX and similar systems for fear that unrestricted data mining could lead to a massive invasion of privacy, as such systems could enable governments to scrutinize the lives and activities of ordinary citizens. Advocates for the MATRIX pilot project counter that this system allows authorized investigators to share and analyze information that is already available to law enforcement from public and state-owned data, without a subpoena or court order. They contend that, with MATRIX, limited investigative information can be developed to generate potential leads within seconds, as opposed to taking days or weeks to manually track and acquire the same information.

The 9/11 Commission did not address the issue of data mining of public or private sector data for the purposes of fighting terrorism, but the commission expressed concern about data sharing between government agencies and the private sector. To protect the privacy of individuals, the commission called for the President to promulgate guidelines to govern information sharing, and establish a board to oversee adherence to those guidelines.

It remains uncertain whether the MATRIX pilot project is currently designed to assess and address privacy and civil liberty concerns. If not, it might be possible that the pilot project could be redesigned to provide an empirical framework to evaluate the use of such data in the future and to minimize unwarranted intrusions on privacy. It has been suggested that, unless a consensus were found regarding the use of public and private sector data for the purposes of national security and counterterrorism, the unregulated use of such data could lead to abuses and unnecessary encroachments on privacy. Perhaps equally as important, a lack of consensus could lead to public rejection and subsequent loss of what many believe to be one of the greatest advantages available to the United States to prevent future terrorist attacks — advanced computing capabilities. This report will be updated as needed.

Contents

The Multi-State Anti-Terrorism Information Exchange (MATRIX) Pilot Project

This report provides an overview of the Multi-State Anti-Terrorism Information Exchange (MATRIX) pilot project, which leverages advanced computer/information management capabilities to more quickly access, share, and analyze *public* records to help law enforcement generate leads, expedite investigations, and possibly prevent terrorist attacks. The pilot project is intended to demonstrate the effective use of such capabilities, but it is less clear whether the project has been designed to prevent unnecessary intrusions on privacy.

The MATRIX pilot project is being administered by the Institute for Intergovernmental Research for the Department of Homeland Security (DHS). Project security and access to the MATRIX system is managed by the Florida Department of Law Enforcement. The project is being funded by the DHS Office of Domestic Preparedness ($8 million) and the Department of Justice (DOJ) Bureau of Justice Assistance ($4 million).

Privacy advocates, civil libertarians, and others oppose MATRIX for fear that such systems will lead to a significant loss of privacy through unrestricted data sharing and mining. Data mining is the use of advanced computer technologies to access data to "discover previously unknown, valid patterns and relationships in large data sets."[1] They oppose data sharing and mining, because in some cases it entails "mass scrutiny of the lives and activities of innocent people," and in their view "constitutes a massive invasion of privacy."[2]

Advocates for the MATRIX pilot project counter that this system allows authorized investigators to share and analyze information that is already available to law enforcement from open public and state-owned data, without a subpoena or court order. They contend that, with MATRIX, limited investigative information can be developed to generate potential leads within seconds, as opposed to taking days or weeks to manually track and acquire the same information and leads.[3] They also

---

[1] Two Crows Corporation, *Introduction to Data Mining and Knowledge Discovery*, third edition (Potomac, MD: Two Crows Corporation, 1999); Pieter Adriaans and Dolf Zantinge, *Data Mining* (New York: Addison Wesley, 1996), as cited in CRS Report RL31798, *Data Mining: An Overview*, by Jeffrey W. Seifert.

[2] American Civil Liberties Union, *The MATRIX: Total Information Awareness Reloaded: New Documents Obtained by ACLU Raise Troubling Questions About Matrix Program*, ACLU Issue Brief #2, May 20, 2004, p. 1, at [http://www.aclu.org/news/NewsPrint.cfm?ID=15834&c=130].

[3] Testimony of Mark Zadra, Florida Department of Law Enforcement, in U.S. Congress, (continued...)

maintain that MATRIX has been developed in compliance with the *National Criminal Intelligence Sharing Plan* that was released by the DOJ in December 2003.[4]

While the 9/11 Commission did not address the issue of aggregating and analyzing public and private sector data (data mining), the commission did express concerns about data sharing between government agencies and the private sector, and called for an executive branch board to oversee, and guidelines to govern, information sharing, so as to protect the privacy of individuals.[5]

Prior to the 9/11 Commission's final report, the *Markle Foundation*[6] *Task Force on National Security in the Information Age* advocated the limited use of both *public* and *private* sector data for national security and counterterrorism purposes, while underscoring concerns about the lack of effective regulation and oversight, and the possible erosion of privacy and civil liberties.[7] In addition, the Department of Defense's (DOD's) Technology and Privacy Advisory Committee raised similar concerns in the wake of that agency's aborted Total Information Awareness program. The committee offered a series of policy and technical recommendations related to the use of data mining in the fight against terrorism that could possibly provide some privacy protection for U.S. persons.[8]

MATRIX Pilot Project

Seisint Inc., a Boca Raton, Florida-based company, developed the advanced computer data storage/retrieval and linking system known as MATRIX.[9] With this

---

[3] (...continued)
House Committee on Government Reform, Subcommittee on Technology, Information Policy, Intergovernmental Relations, and Census (Washington, July 13, 2004), p. 2.

[4] U.S. Department of Justice, Office of Justice Programs, *The National Criminal Intelligence Sharing Program: Solutions and Approaches for a Cohesive Plan to Improve our Nation's Ability to Share Criminal Intelligence*, Oct. 2003, 83 pp.

[5] National Commission on Terrorist Attacks Upon the United States, *The 9/11 Commission Report: Final Report of the National Commission on Terrorist Attacks upon the United States* (Washington: GPO, 2004), pp. 394-395.

[6] The John and Mary R. Markle Foundation was established in 1927 "to promote the advancement and diffusion of knowledge ... and the general good of mankind." In 1998, under Zoe Baird's leadership, the Foundation focused its efforts on addressing critical public needs in the information age. As part of its National Security program, the Foundation has examined how best to mobilize information and information technology to improve national security while protecting civil liberties.

[7] Markle Foundation, Task Force on National Security in the Information Age, *Creating a Trusted Network for Homeland Security*, Second Report, 2003, p. 4.

[8] U.S. Department of Defense, *Safeguarding Privacy in the Fight Against Terrorism: Report of the Technology and Privacy Advisory Committee*, Mar. 2004, 119 pp.

[9] Seisint was recently acquired for $775 million by an Anglo/Dutch group — Reed Elsevier, the same company that owns Lexis-Nexis. See Chris Nuttal, "Big Brother's Little Brother' Comes under Reed's Control: The Purchase of Information Specialist Seisint Will Give (continued...)

system, recent press accounts state that Seisint compiled a list of 120,000 names of persons who were identified as having "high terrorism factor (HTF)" scores in the weeks immediately following the September 11, 2001, terrorist attacks. Without identifying information provided by federal investigators, the system reportedly identified five of the September 11, 2001, highjackers.[10] The list was provided to the Federal Bureau of Investigation (FBI), Immigration and Naturalization Service (INS), U.S. Secret Service, and the Florida Department of Law Enforcement (FDLE).[11] According to a Seisint presentation acquired by the American Civil Liberties Union (ACLU), many arrests resulted from this list.[12]

The *Washington Post* reported that in January 2003, Seisint founder Hank Asher[13] — accompanied by Florida Governor Jeb Bush and then FDLE Commissioner Tim Moore — briefed Vice President Dick Cheney, Homeland Security Secretary Thomas Ridge, and FBI Director Robert Mueller on MATRIX. Soon thereafter, federal funding was allocated for the MATRIX pilot project.[14]

The Institute for Intergovernmental Research (IIR)[15] is administering the MATRIX pilot project for the DHS and the DOJ. According to the MATRIX website, IIR is preparing a "proof of concept" study for these departments.[16] The MATRIX pilot project is reportedly funded by the DHS Office of Domestic Preparedness (ODP) — $8 million, and the DOJ Bureau of Justice Assistance (BJA)

---

[9] (...continued)
Anglo/Dutch Company an Unusual Insight into the World," *Financial Times* (London, England, July 15, 2004), p. 22.

[10] The Associated Press, "Database Firm Gave Feds Terror Suspects: 'Matrix' Developer Turned Over 120,000 Names," *MSNBC*, May 20, 2004, at [http://www.msnbc.msn.com/id/5020795/].

[11] American Civil Liberties Union, *The MATRIX: Total Information Awareness Reloaded*, ACLU Issue Brief #2, May 20, 2004, p. 2.

[12] American Civil Liberties Union, "ACLU Unveils Disturbing New Revelations About MATRIX Surveillance Program," press release, May 20, 2004, available at [http://www.aclu.org/news/NewsPrint.cfm?ID=15834&c=130].

[13] Asher reportedly resigned from Seisint Inc. when questions about his background arose during a Matrix-related $1.6 million contract negotiation. It was alleged that Asher was an informant for state and federal law enforcement in a cocaine smuggling investigation in which he may have piloted an aircraft used to smuggle the narcotic into the United States. See Cynthia L. Webb, "Total Information Dilemma," *Washington Post*, May 27, 2004.

[14] Robert O'Harrow, Jr., "Anti-Terror Database Got Show at White House," *Washington Post*, May 21, 2004, p. A12.

[15] The IIR is a Florida-based research and training organization that specializes in nonpartisan and inter-governmental studies of law enforcement, juvenile justice, and criminal justice issues. See [http://www.matrix-at.org/contact_matrix.htm] for MATRIX contacts.

[16] See [http://www.matrix-at.org/roles.htm].

— $4 million.[17] MATRIX utilizes the Regional Information Sharing System (RISS).[18] Composed of several secure regional intranets, RISS is a multi-state and multi-agency system that was originally established with BJA assistance as a means of sharing drug enforcement-related criminal intelligence.[19]

Seisint executive Bill Shrewsbury reportedly has asserted that the MATRIX system no longer uses the scoring system that produced the "high terrorist factor" list of suspects in the weeks following the September 11, 2001, attacks. Among other things, Seisint apparently no longer has access to "intelligence data" that were previously fed into the system for demonstration purposes.[20]

The core of the MATRIX pilot project is the Factual Analysis Criminal Threat Solution (FACTS). With this new application, law enforcement officers can more effectively query available public records with incomplete information — such as a partial license plate number.[21] In addition to the partial vehicle tag query, FACTS includes crime mapping, association charting (described below), and lineup and photo montage applications. There are over 3.9 billion public records available in the FACTS database.[22]

The FDLE controls access to the MATRIX pilot program and is charged with the responsibility for assuring that the FACTS application is secure. To participate in the MATRIX pilot project, states sign an agreement that the system will only be used in the pursuit of criminal investigative matters. According to the MATRIX website, information sharing between the states is arranged so that each state is in compliance with its privacy laws.[23]

While five states are currently participating in the MATRIX pilot project, 11 states have dropped out, prompted by concerns in part about how the data were being stored in a central repository. According to press accounts, to participate in the system, states initially were required to transfer control of state-owned data to MATRIX system administrators.[24] The transfers of these data to outside concerns may have violated state privacy laws. The states of Connecticut, Florida, Michigan, Ohio, and Pennsylvania are currently participating in the MATRIX pilot project. The states of Alabama, California, Georgia, Kentucky, Louisiana, New York, Oregon,

---

[17] Blake Harrison, "MATRIX Revolution: Sophisticated Technology Allows Law Enforcement Across the Nation to Communicate Quickly and Solve Crimes," *State Legislatures*, May 2004, p. 14.

[18] William Welsh, "Matrix Taps Databases," *Washington Technology*, Sept. 1, 2003.

[19] For more information on RISS, see [http://www.iir.com/riss/].

[20] Cynthia L. Webb, "Total Information Dilemma," *Washington Post*, May 27, 2004. See [http://www.washingtonpost.com/ac2/wp-dyn/A60986-2004May27?language=printer].

[21] See [http://www.matrix-at.org/FACTS_defined.htm].

[22] See [http://www.matrix-at.org/newsletter.pdf].

[23] See [http://www.matrix-at.org/matrix_defined.htm].

[24] Justin Rood, "Controversial Data-Mining Project Finds Ways Around Privacy Laws," *CQ Homeland Security — Intelligence*, July 23, 2004, p. 1.

South Carolina, Texas, Utah and Wisconsin have dropped out of the project. The original 16 participating states would have covered over half of the U.S. population.

To alleviate some of the privacy concerns, the developers of MATRIX re-engineered the system on the principle of distributed computing. Rather than housing all the data in a central repository, the re-engineered MATRIX system accesses data remotely by using web-based and secured connectivity to individual state-maintained databases.[25] Under this arrangement, if the pilot program receives no further funding, the data remain with the states.

In its report entitled *Creating a Trusted Network for Homeland Security*, the Markle Foundation's Task Force on National Security in the Information Age has suggested that unless a consensus were found about the use of public and private sector data for criminal intelligence purposes, efforts like MATRIX run the risk of being cut back or eliminated.[26] A case in point would be the Total Information Awareness (TIA) program, a DOD program that generated controversy in 2002, when it became known that the Defense Advanced Research Projects Agency (DARPA) had created a tool to "find new sources of data, mine information from the new and old, generate information, make it available for analysis, convert it to knowledge, and create actionable options."[27] Fears that TIA would be used against U.S. persons (citizens or legal permanent residents) prompted Congress to restrict the funding for DOD data mining, so it could only be used for "processing, analysis, and collaboration tools" for counterterrorism with a *foreign* nexus.[28]

Some privacy advocates have compared MATRIX to the TIA program.[29] While the MATRIX pilot project is on a smaller scale than the TIA program, both programs are affected by many of the same laws and underlying privacy issues.[30] The FDLE, however, maintains that MATRIX is not similar to TIA.[31] MATRIX queries only public and state-owned records (state vehicle and crime records), as opposed to private records (mailing lists and credit histories). The FDLE underscores that these

---

[25] Ibid., p. 1.

[26] Markle Foundation, *Creating a Trusted Network for Homeland Security*, p. 4.

[27] John Poindexter, Overview of the Information Awareness Office, prepared remarks for delivery at DARPATech 2002, Anaheim, CA, Aug. 2, 2002, at 1, as cited in *Safeguarding Privacy in the Fight Against Terrorism: The Report of the Technology and Privacy Advisory Committee*, Department of Defense, Mar. 2004, p. vii.

[28] Ibid, p. vii.

[29] American Civil Liberties Union, "ACLU Unveils Disturbing New Revelations About MATRIX Surveillance Program," press release, May 20, 2004, at [http://www.aclu.org/Privacy/Privacy.cfm?ID=15834&c=130].

[30] For further information about privacy laws and the TIA program, see CRS Report RL31730, *Privacy: Total Information Awareness Programs and Related Information Access, Collection, and Protection Laws*, by Gina Marie Stevens.

[31] Testimony of Mark Zadra Member, Florida Department of Law Enforcement, in U.S. Congress, House Committee on Government Reform, Subcommittee on Technology, Information Policy, Intergovernmental Relations, and Census (Washington, July 13, 2004), p. 5.

public and state-owned records accessed by MATRIX are available to law enforcement without a subpoena or court order.[32]

The FDLE also maintains that MATRIX — in its most current iteration — has been developed in compliance with the *National Criminal Intelligence Sharing Plan* that was released by the DOJ in December 2003.[33]

Matrix Data Sources

As currently configured, the MATRIX system accesses a broad array of public data, ranging from motor vehicle driving records to bankruptcy filings. While much of these data have been available to law enforcement, they have not been previously queried, cross-referenced, and analyzed with computers. According to the MATRIX website, such records include the following:

- pilot licenses issued by the Federal Aviation Administration;
- aircraft ownership;
- property ownership;
- U.S. Coast Guard-registered vessels;
- state sexual offender lists;
- corporate filings;
- Uniform Commercial Code filings or business liens;
- bankruptcy filings; and
- state-issued professional licenses.[34]

According to the MATRIX website, available records also include records that have historically been available to law enforcement agencies. Such records include

- criminal history records;
- department of corrections information and photo images;
- sexual offender criminal history files;
- driver's license information and photo images; and
- motor vehicle registration information.[35]

According to the MATRIX website, such records *do not* include the following:

- telemarketing calling lists;
- direct mail mailing lists;
- airline reservations or travel records;
- frequent flyer/hotel stay program membership or activity;
- magazine subscription lists or reading lists;

---

[32] Ibid., p. 2.

[33] U.S. Department of Justice, Office of Justice Programs, *The National Criminal Intelligence Sharing Program: Solutions and Approaches for a Cohesive Plan to Improve our Nation's Ability to Share Criminal Intelligence*, Oct. 2003, 83 pp.

[34] See [http://www.matrix-at.org/data_sources.htm].

[35] Ibid.

- telephone logs or records;
- credit card or debit numbers;
- purchases (from most sources);
- mortgage or car payments;
- bank account numbers or account balances;
- the costs of a home addition;
- birth certificates;
- marriage licenses;
- divorce decrees; or
- utility bill payments.[36]

The *Washington Post* has quoted Guy Tunnell, the current FDLE Commissioner and sitting chair of the MATRIX executive committee, stating that the use of this system will be driven by threat information developed from criminal and domestic security intelligence, or by actual criminal investigations. He argues that MATRIX is an investigative tool, which will not be used indiscriminately to conduct surveillance on individuals, and that the system is designed to guard against inappropriate or unauthorized use.[37]

Privacy Concerns

Privacy advocates, civil libertarians, and others object to the MATRIX pilot project, because it could give the government greater and easier access to vast amounts of both public, and possibly private sector, data that are gathered — oftentimes with expectations of privacy — for purposes other than law enforcement. They are also concerned with the quality of these data, and the ability of individuals to correct faulty or invalid data about themselves — a prospect that would require some transparency in regard to the system and its data sources. This would be especially important for any system that would be used to screen persons and may result in the denial of a benefit, the ability to travel, or some other activity.

Privacy advocates, civil libertarians, and others reportedly object to the possibility that the government could launch initiatives with the aid of computer programs (algorithms) that could be used to electronically search for, or monitor, patterns of suspicious behavior of persons whom the government previously had no reason to suspect of criminal or terrorist activities.[38] In addition, they are concerned about the characteristics (such as, life style, ethnicity, religious faith, or income status) with which the government may choose to profile individuals under certain circumstances with this system.

---

[36] Ibid.

[37] Cynthia L. Webb, "Total Information Dilemma," *Washington Post*, May 27, 2004. See [http://www.washingtonpost.com/ac2/wp-dyn/A60986-2004May27?language=printer].

[38] Robert O'Harrow, Jr., "U.S. Backs Florida's New Counterterrorism Database: 'Matrix' Offers Law Agencies Faster Access to Americans' Personal Records," *Washington Post*, Aug. 6, 2003, p. A01.

Proponents argue that such capabilities could be of immense value to intelligence and law enforcement authorities charged with national security, counterterrorism, and law enforcement responsibilities. For example, some state that such capabilities could be used to provide valuable leads for child abductions or to identify the whereabouts of known Al-Qaeda operatives. While many may agree with these uses under certain circumstances, there may also be concern about "mission creep,"[39] as these same capabilities could be used to monitor the activities of social activists and other dissenters who are opposed to governmental policies. Some argue, such a system could be used for political purposes, such as undermining one's political opponents.

Policy Considerations Related to Data Mining for Counterterrorism Purposes

According to the *Markle Foundation Task Force on National Security in the Information Age*, in the past decade, the quantity of personal data held by the private sector has exploded, as computing and storage capabilities have rapidly advanced, and associated costs have correspondingly diminished.[40] The same could be said of public data held by federal, state, and local governments.

Much public and private sector data have been aggregated into "data marts." This information is often available commercially for sale from companies specializing in data aggregation, like ChoicePoint, Equifax, Experian, Qsent, LexisNexis, and Westlaw. With advanced computing technologies tera- and peta-bytes of data can be manipulated,[41] and multiple data marts can be merged or cross-referenced. Moreover, computer applications are available to "mine" these data for the purposes of profiling, pattern analysis, link analysis,[42] transactional footprinting,[43] and identity verification.

---

[39] Markle Foundation, *Creating a Trusted Network for Homeland Security*, p. 31.

[40] Ibid., p. 30.

[41] A terabyte is roughly a trillion bytes (10^{12}). A petabyte is roughly a quadrillion bytes (10^{15}).

[42] In a law enforcement and intelligence context, "link analysis" means uncovering relationships that may be indicative of suspicious patterns, groups, or connections. Oftentimes, these relationships are diagramed to facilitate further analysis. Sophisticated link analysis programs are often capable of calculating the statistical significance of the diagramed relationships. For further information on link analysis, see Hamid R. Nemati and Christopher D. Barko, *Organizational Data Mining: Leveraging Enterprise Data Resources for Optimal Performance* (Idea Group Publishing, 2004), p. 145.

[43] In the records of online electronic commerce and other Internet activity, investigators can glean the data trails of suspicious activities by individuals and groups. This is known as "transactional footprinting." Systematic analysis of such data has been useful in cases involving identity theft, credit card fraud, health care fraud, cyberstalking, and web scams. For an example of "transactional footprinting," see Associated Press, "MSU Professor Helps Track Terrorists: FBI Will Use Process She Created To Search Online for Identity Thieves," *Detroit News*, Mar. 20, 2003, at [http://www/detnews.com/2003/metro/0303/20/d12e-112610.htm].

Even before the September 11, 2001, terrorist attacks, law enforcement agents in the United States had availed themselves of the Seisint's services for the purposes of furthering criminal investigations under limited circumstances. For example, through DNA analysis, the FBI national data center had linked several sexual assaults — one involving a murder — in Philadelphia, Pennsylvania and Fort Collins, Colorado. The Philadelphia police turned to Seisint to produce a list of 40 men who had lived in both cities at the time of the attacks. While clearly this list included persons who were innocent of the crimes under investigation, the rapist reportedly was identified from this list of potential suspects. He was convicted of murder in Colorado and is serving a life sentence.[44]

While, on the face of it, few would condemn the capture and successful prosecution of a rapist and murderer, privacy advocates, civil libertarians, and others have questioned whether appropriate policies and safeguards have been established to prevent the abuse of such information. They maintain that unrestricted data mining could lead to a massive invasion of privacy, as such systems could enable governments to scrutinize the lives and activities of ordinary citizens.[45] At the same time, public and private sector data can be used to help identify known and suspected terrorists — as well as violent criminals. Such efforts are already underway at the FBI-administered Foreign Terrorist Tracking Task Force (FTTTF).[46] The Government Accountability Office (formerly the General Accounting Office) recently surveyed 128 federal agencies, and reported that 52 agencies are currently engaged in 199 data mining projects for a wide range of purposes, but only a few involve analyzing foreign and criminal intelligence to detect terrorist and criminal activities.[47]

More recently, the 9/11 Commission issued its final report in July 2004. While the commission did not address the issue of aggregating and analyzing public and private sector data (data mining), the commission did expressed concerns about data sharing. The commission called for guidelines — to be determined by the President — that would govern information sharing between government agencies and the private sector, so as to protect the privacy of individuals about whom the information was being shared.[48] The commission also called for the establishment of an executive branch board to oversee adherence to the guidelines.[49]

---

[44] Blake Harrison, "MATRIX Revolution," *State Legislatures*, May 2004, p. 13.

[45] American Civil Liberties Union, *The MATRIX: Total Information Awareness Reloaded: New Documents Obtained by ACLU Raise Troubling Questions About Matrix Program*, ACLU Issue Brief #2, May 20, 2004, p. 1, at [http://www.aclu.org/news/NewsPrint.cfm?ID=15834&c=130].

[46] For further information, see CRS Report RL32366, *Terrorist Identification, Screening, and Tracking Under Homeland Security Presidential Directive 6*, by William J. Krouse.

[47] U.S. Government Accountability Office, *Data Mining: Federal Efforts Cover a Wide Range of Uses*, GAO-040548, May 2004, p. 2.

[48] National Commission on Terrorist Attacks upon the United States, *The 9/11 Commission Report*, p. 394.

[49] Ibid., p. 395.

Prior to the 9/11 Commission's final report, the Markle task force reported that there remain no government-wide policy guidelines for the acquisition, storage, and retention of privately-held data by the federal government for counterterrorism, national security, or other law enforcement purposes. The report included a recommendation that new guidelines should be promulgated by the executive branch that would address the following issues: (1) government acquisition and use of private sector data; (2) government retention of the data; (3) sharing of the data by the acquiring agency with other agencies for purposes other than counterterrorism; and (4) accountability and oversight.[50] This report also recommended that Congress provide strong oversight of government-wide activities related to the acquisition, retention, and sharing of private sector data for the purposes of national security, counterterrorism, and law enforcement.

In addition, DOD's Technology and Privacy Advisory Committee (TAPAC), which was formed in the wake of the TIA controversy, issued a report that concluded that data mining could be a vital tool in the fight against terrorism, but it presented significant privacy concerns, when such endeavors involved data on U.S. persons[51] The TAPAC called for a statutory and regulatory framework for data mining that would, among other things, require written findings by an agency head and court authorization before data mining could be conducted if it involved U.S. persons.[52] Other technical requirements recommended by the TAPAC included data minimization and anonymization,[53] secured systems, and immutable audit trails.[54]

As stated earlier in this report, it remains uncertain whether the MATRIX pilot project is currently designed to assess and address privacy and civil liberty concerns. If not, it might be possible that the pilot project could be redesigned to provide an empirical framework to evaluate the use of such data in the future and to minimize unwarranted intrusions on privacy. The Markle report suggested that, unless a consensus were found regarding the use of public and private sector data for the purposes of national security and counterterrorism, the unregulated use of such data could lead to abuses and unnecessary encroachments on privacy. Perhaps equally as important, a lack of consensus could lead to public rejection and subsequent loss of what many believe to be one of the greatest advantages available to the United States to prevent future terrorist attacks — advanced computing capabilities.

---

[50] Markle Foundation, *Creating a Trusted Network for Homeland Security*, p. 33.

[51] U.S. Department of Defense, *Safeguarding Privacy in the Fight Against Terrorism: Report of the Technology and Privacy Advisory Committee*, Mar. 2004, p. viii.

[52] Ibid., pp. 49-50.

[53] *Data minimization* means the data mining system should access, disseminate, and retain the least amount of data consistent with the data mining mission. *Data anonymization* entails removing, encrypting, or otherwise obscuring information by which specific individuals can be identified, such as, addresses, phone numbers, and social security numbers. As part of this process, persons identified through data mining as possibly connected with terrorist activities or other crimes would have to be reidentified.

[54] Ibid., p. 50.